



Maori Ito @ NIBIO

**Enhanced search for life science  
databases with proposed  
schema.org extension**







Oatmeal Cookies  
3/4 cup shortening  
1 cup brown sugar  
1/2 cup sugar  
1 egg  
1/4 cup water  
1 tsp vanilla  
1 cup flour  
1 tsp salt  
1/2 tsp soda  
5 cups oats  
Cream shortening, sugars, eggs, water  
and vanilla. Add dry ingredients,  
mix well. Stir in oats. Drop by

Toll House Cookies  
3/4 cup shortening  
1 cup brown sugar  
1/2 cup sugar  
1 egg  
1/4 cup water  
1 tsp vanilla  
1 cup flour  
1 tsp salt  
1/2 tsp soda  
5 cups oats  
Cream shortening, sugars, eggs, water  
and vanilla. Add dry ingredients,  
mix well. Stir in oats. Drop by

E.g. secret recipe





A very healthy, spicy and tangy  
rice noodle salad is one dish from  
Singapore

**Ingredients**

- 200gms chicken (minced) soaked in 300ml oil water
- 200gms chicken
- 2 big onions
- 100gms blended dried chilies
- 1/2 cup lime juice
- 200gms sugar
- 200gms rice flour (add 100gms rice flour)
- 100gms chicken (dried) (dried)

**Method**

1. Blend chicken & chicken. Fry till dry. (It is important that this step is not hurried. If the chicken is not fried well enough the dish will have a 'wet' taste.)
2. Squeeze chicken and strain for juice. Set aside.
3. Fry chicken in oil (1/2 above) separately. Add in chicken (10 above) low heat (2 above) and fried chicken (2 above).
4. Add chicken water (1 above), sugar (2 above) and bring to boil. Leave to simmer.

**Ingredients for Serving**

- 200gms chicken (thin rice noodles)
- 200gms low fat (dried) (dried)
- 100gms low fat (dried) (dried)
- 100gms low fat (dried) (dried)
- 100gms low fat (dried) (dried)
- 100gms low fat (dried) (dried)



# Why Cross Search?

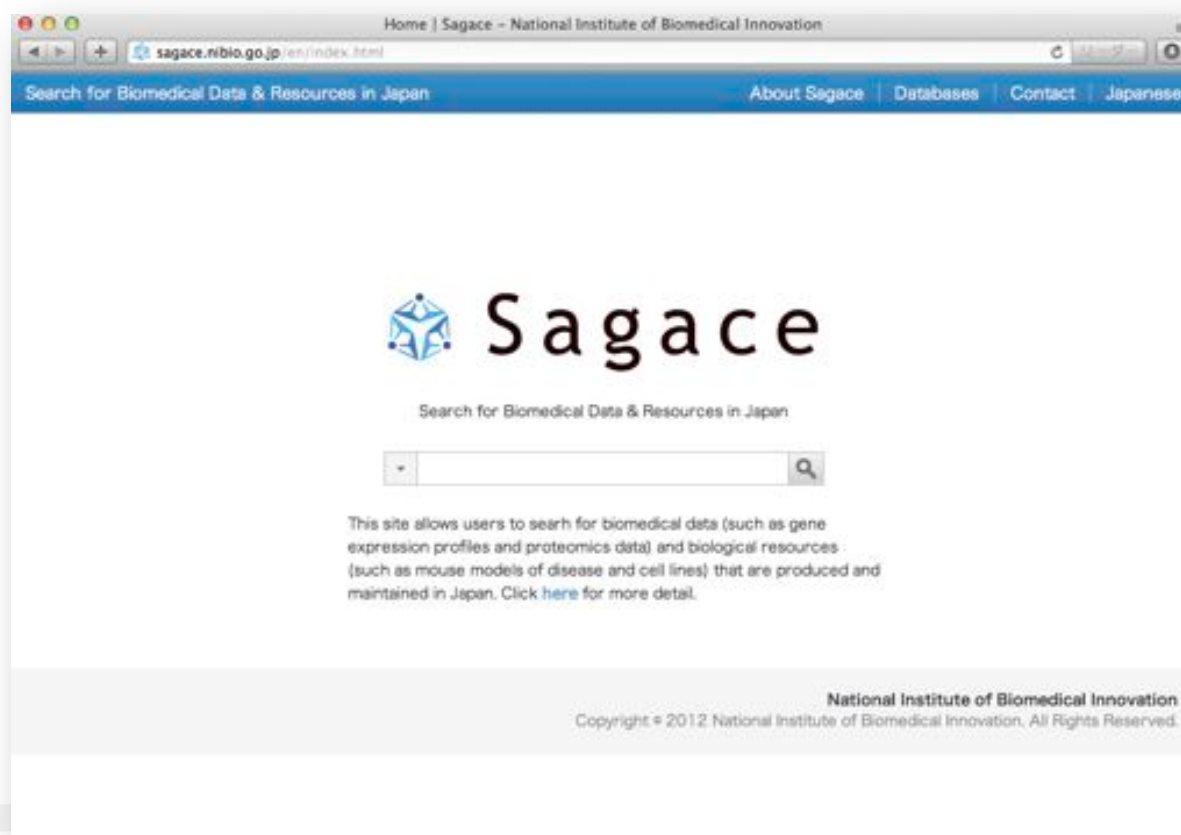
- Easy to use
- Accustomed to use
- Appropriate for comparing various kinds of databases
- Major cross engine services will continue for a while.



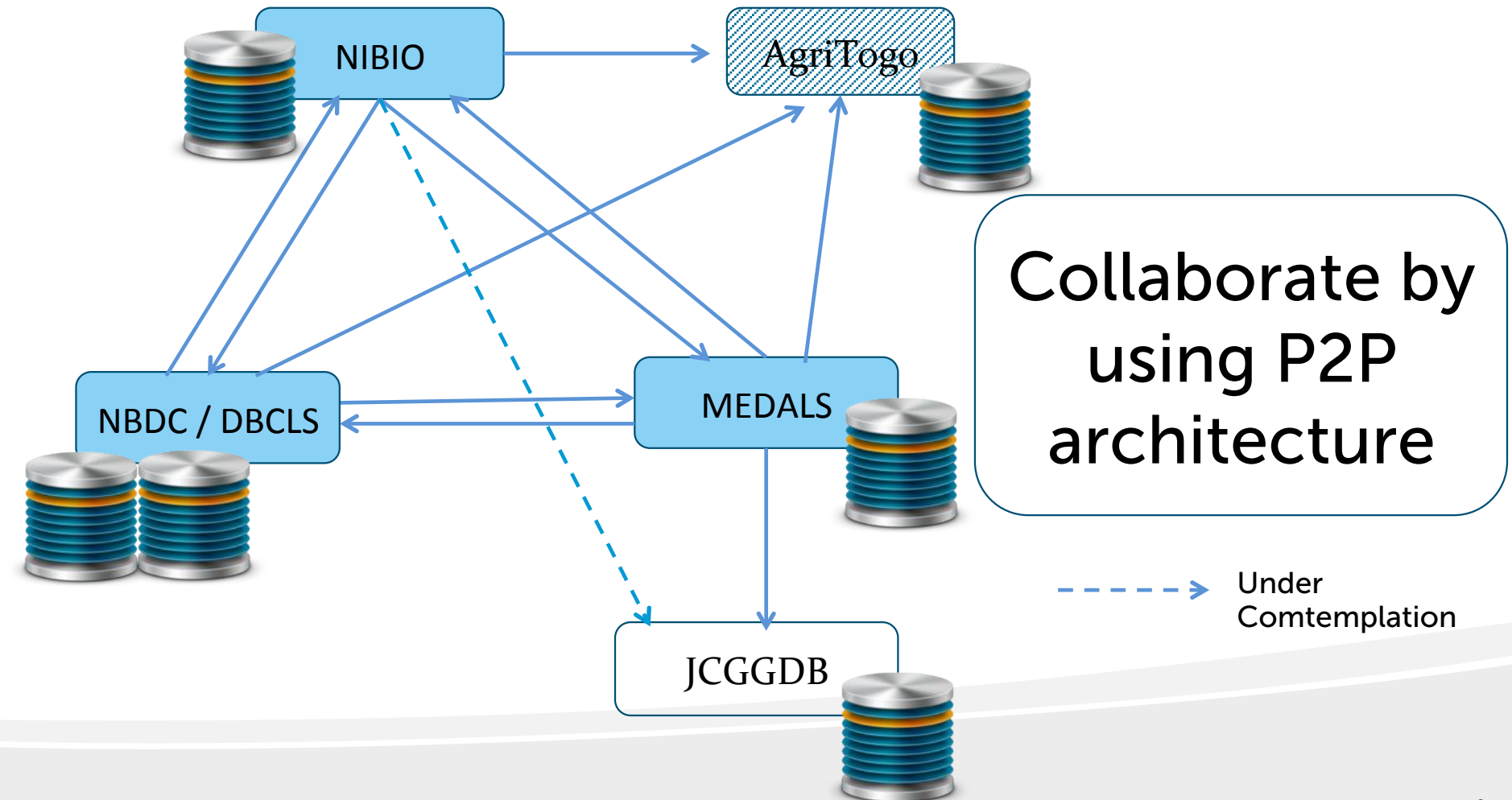


# Sagace

- Search for Biomedical Data & Resources in Japan



# Search System





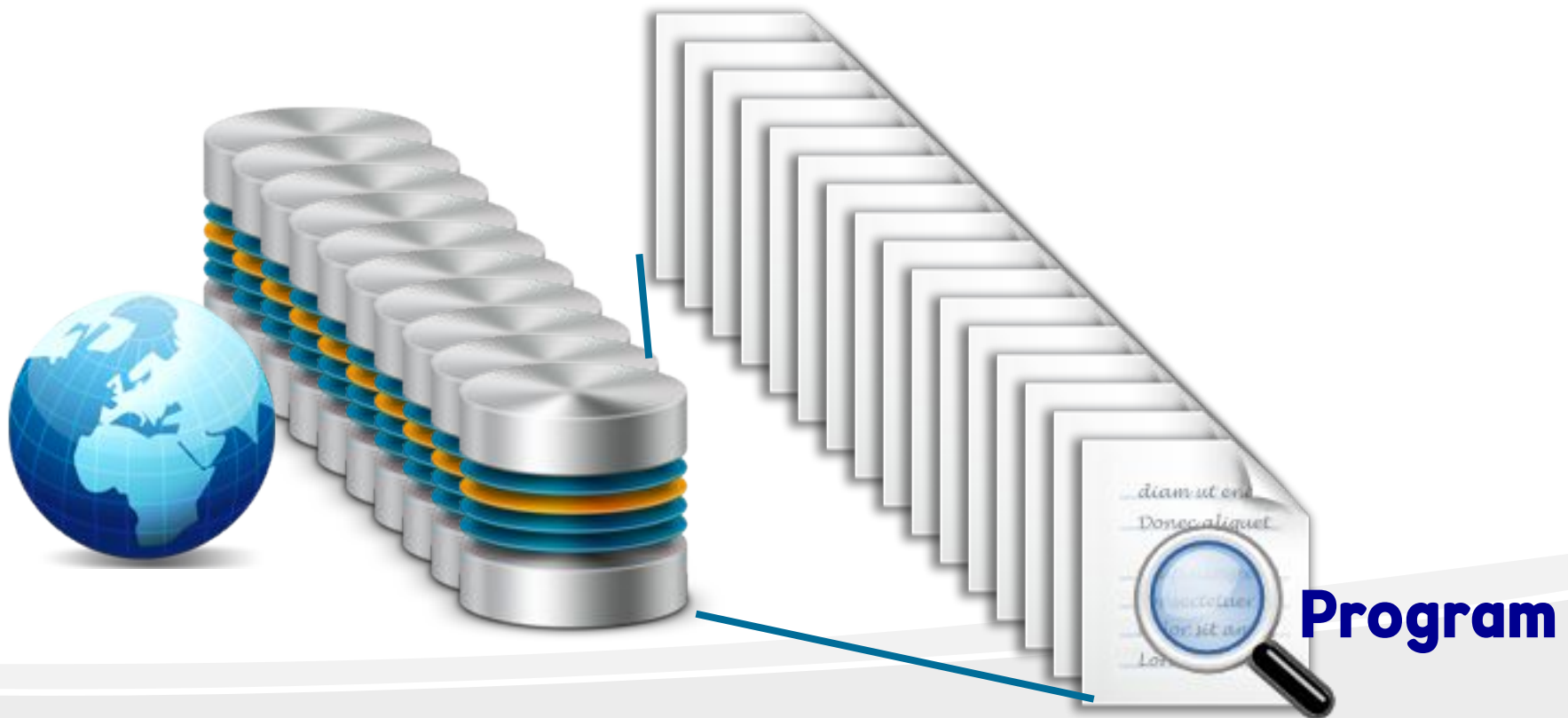


# Mechanism of Search Engine

1. Crawling
2. Indexing
3. Query Processing
4. Scoring

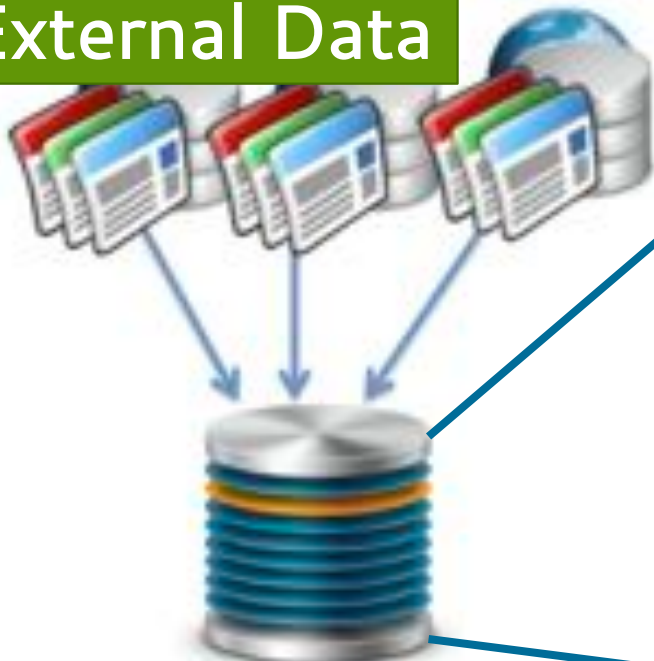
# Crawling

- Crawl databases and pages by program



- Split data convenient size and store own server

## External Data



# Internal Server

```
@uri=http://cellbank.nibio.go.jp/celldata/icrb0527.htm  
@uri=http://cellbank.nibio.go.jp/celldata/icrb0527.htm  
@uri=http://cellbank.nibio.go.jp/celldata/icrb0527.htm  
@uri=http://cellbank.nibio.go.jp/celldata/icrb0527.htm  
@uri=http://cellbank.nibio.go.jp/celldata/icrb0527.htm  
@uri=http://cellbank.nibio.go.jp/celldata/icrb0527.htm  
@uri=http://cellbank.nibio.go.jp/celldata/icrb0527.htm  
@title=TIG-102 - JCRB細胞バンク  
@db=cellbank  
@entryID=JCRB0527  
@species=Human  
@taxonID=9606  
  
JCRB0527 TIG-102 human Homo sapiens F 97 year-old  
ploidy fibroblast fibroblast-like skin, normal diploid  
K. Japanese Normal human skin fibroblast cell line  
normal diploid (2n=46, XX)
```

# Query Processing and Scoring



All  
Biological Resources (79)  
Patents (43)  
Plants (24)  
Microbes (33)  
None specified (45)

Search Results: 79 hits Option:  Selected Facet Item[s]: Biological Resources

## TIG-102 - JCRB細胞バンク

<http://cellbank.nibio.go.jp/> JCRB細胞バンク

Biological Resources | Human, Animals (non-human) | Cell & Tissue - ID: JCRB0527 -

Species: Human

JCRB0527 TIG-102 human Homo sapiens F 97 year-old skin normal skin, normal diploid fibroblast fibroblast-like skin, normal diploid fibroblast normal Yamamoto, K. Japanese Normal human skin fibroblast cell line. Temporary no stock. skin normal diploid (2n=46, XX)

## TIG-101 - JCRB細胞バンク

<http://cellbank.nibio.go.jp/> JCRB細胞バンク

Biological Resources | Human, Animals (non-human) | Cell & Tissue - ID: JCRB0526 -

Species: Human

JCRB0526 TIG-101 human Homo sapiens skin normal skin, normal diploid fibroblast fibroblast-like skin, normal diploid fibroblast Yamamoto, K. Japanese Normal human diploid cell from skin

## XP24KO - JCRB細胞バンク

<http://cellbank.nibio.go.jp/> JCRB細胞バンク

Biological Resources | Human, Animals (non-human) | Cell & Tissue - ID: JCRB3012 -

Species: Human

JCRB3012 XP24KO human Homo sapiens F skin Xeroderma pigmentosum (E) fibroblast-like skin fibroblast primary Hereditary disease with high cancer risk xeroderma pigmentosum (autosomal recessive) complementation group E skin fibroblast, skin

@uri=http://cellbank.nibio.go.jp/celldata/jcrb0527.htm  
@title=TIG-102 - JCRB細胞バンク  
@db=cellbank  
@entryID=JCRB0527  
@species=Human  
@taxonID=9606

JCRB0527 TIG-102 human Homo sapiens F 97 year-old skin normal skin, normal diploid fibroblast fibroblast-like skin, normal diploid fibroblast normal Yamamoto, K. Japanese Normal human skin fibroblast cell line. Temporary no stock. skin normal diploid (2n=46, XX)

@uri=http://cellbank.nibio.go.jp/celldata/jcrb0526.htm  
@title=TIG-101 - JCRB細胞バンク  
@db=cellbank  
@entryID=JCRB0526  
@species=Human  
@taxonID=9606

JCRB0526 TIG-101 human Homo sapiens skin normal skin, normal diploid fibroblast fibroblast-like skin, normal diploid fibroblast Yamamoto, K. Japanese Normal human diploid cell from skin

@uri=http://cellbank.nibio.go.jp/celldata/jcrb3012.htm  
@title=XP24KO - JCRB細胞バンク  
@db=cellbank  
@entryID=JCRB3012  
@species=Human  
@taxonID=9606

JCRB3012 XP24KO human Homo sapiens F skin Xeroderma pigmentosum (E) fibroblast-like skin fibroblast primary Hereditary disease with high cancer risk xeroderma pigmentosum (autosomal recessive) complementation group E skin fibroblast, skin



# ~~Bad~~ Skeptical Reputations for Search Results...

- Why?
- Cannot get useful information immediately.




# How to improve accuracy?

- What is accuracy for life science database cross search?
- What is accuracy for life science specialist?





- 
- In general, developers emphasize search algorithms and scorings.
  - However, general results and methods for cross search may not be suitable for life science specialists..?
  - Data (Index files) from life science databases are sometimes difficult to understand immediately.
  - It's hard to make each crawler program for each database and maintainance it.
  - (We have no extra .... to make proper search page like entrez et al....)

# Metadata!

- One way to solve these problems

Difficult to  
understand  
data  
immediately





If metadata are added data...

Data



Metadata

Disease:Epithelial adenoma  
Species:Mouse  
Keywords:DNA sequence  
Last Modified:2013-01-19

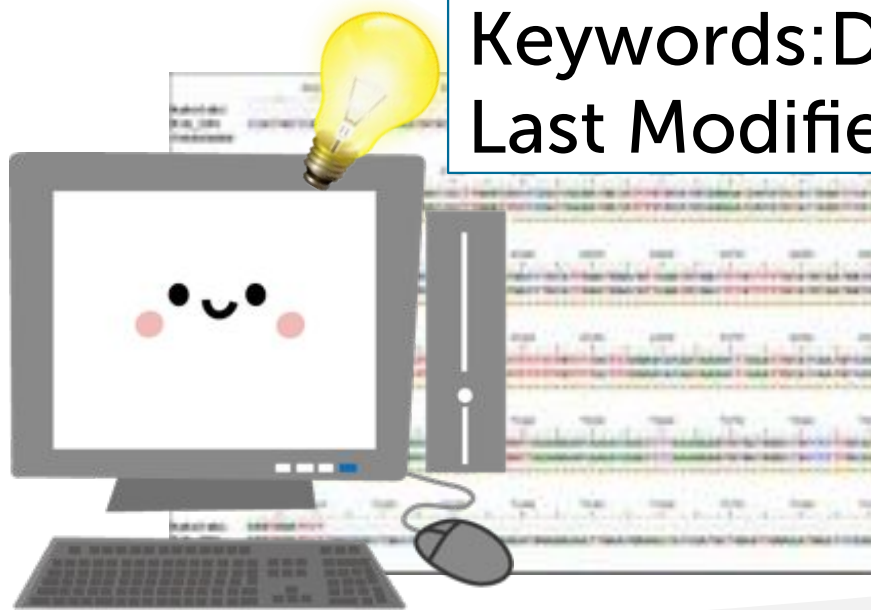




# Easy to pick up for crawlers

## Metadata

Disease:Epithelial adenoma  
Species:Mouse  
Keywords:DNA sequence  
Last Modified:2013-01-19



# Easy to understand for users

- It can be a guide to improve user experience.

EMDB-1556: RsbR · RsbS · RsbT複合体 (ストレスゾーム) 「単粒子」状態 /

...

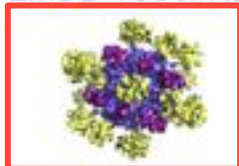
<http://www.pdbj.org/> 3D Electron Microscopy (3D-EM) Data Navigator

その他のデータベース | ヒト, 動物 (ヒト以外), 植物, 微生物 | タンパク質, 細胞・組織

EM Navigator EMD-1556 RsbR · RsbS · RsbT複合体 (ストレスゾーム) 「単粒子」状態 /

8.3 Å分解能 エンe', a signal transduction hub 試料名・マップの注釈 RsbR · RsbS · RsbT複合体 (ストレスゾーム) Ternary RsbR Rsb .....

EMDB-1556: RsbR · RsbS · RsbT複合体 (ストレスゾーム) 「単粒子」状態



<http://www.pdbj.org/>

3D Electron Microscopy (3D-EM) Data Navigator

その他のデータベース | ヒト, 動物 (ヒト以外), 植物, 微生物 | タンパク質

細胞・組織 - ID: 1556, - Species: Bacillus subtilis, - Last Modified: 2012-10-24

EM Navigator EMD-1556 RsbR · RsbS · RsbT複合体 (ストレスゾーム) 「単粒子」状態 / 8.3 Å分解能 エン合体 (ストレスゾーム) Ternary RsbR RsbS RsbT complex 著者 Jon Marles-Wright, Tim Grant, Olivier Delumeau .....

Image



# How to use it?

- Mark up data by microdata like a tag

The screenshot shows the EMDB entry page for ID 1556. The page is divided into several sections, each with a microdata tag overlaid:

- Image**: A blue square tag with a string, pointing to the EMDB logo in the top left corner.
- Title**: A purple square tag with a string, pointing to the title "Molecular Architecture of the 'stressosome', a signal transduction hub".
- ID**: A purple square tag with a string, pointing to the ID "1556".
- Last Modified**: A blue square tag with a string, pointing to the "Last update" date "2012-10-24".

The page content includes:

- Database / ID**: EM DATA BANK (EMDB) / 1556
- Title**: Molecular Architecture of the 'stressosome', a signal transduction hub
- Map**: Ternary Raft RaG RaG complex
- Sample**: Ternary Raft RaG RaG complex
- Keywords**: [Raft](#), [RaG](#), [Stressosome](#), [Raft](#), [RaG](#), [Stressosome](#), [Raft](#)
- Authors**: [Marta Wozniak](#), [Grant T. Delmonico](#), [G. van Duinen](#), [G. Tibbels](#), [S. J. Lewis](#), [P. J. Morris](#), [J. W. Noorman](#), [J. A. Quin](#), [M. B. Saxe](#), [F. S. Babou](#), [A. Tschopp](#), [W. van Hees](#), [M. Lewis](#), [R. J.](#)
- Date**: Deposition: 2008-09-08, Header release: 2008-09-11, Map release: 2009-04-16, Last update: 2012-10-24
- EMDB Sites**: [EMDB](#), [RCSB](#), [PDB](#), [EMBL](#), [EMDB](#), [RCSB](#), [PDB](#)
- Structure Visualization**: [Play](#), [Medium](#), [Large](#), [X Off](#), [More Page](#)
- Movies**: [#1](#), [#2](#)
  - [#1](#): Surface view with section colored by density value, Surface level: 24, Made by [UCSF Chimera](#)
  - [#2](#): Surface view colored by radius, Surface level: 24, Made by [UCSF Chimera](#)
- Supplemental Images**: [emol328.png](#)
- Structure viewers**: [VMD](#), [Launch PyMOL](#), [About PyMOL](#), [Volume Viewer](#), [RCSB PDB](#)
- Related Structure Data**: [EMDB 1552](#), [EMDB 1553](#), [EMDB 1554](#), [EMDB 1555](#), [EMDB 1556](#), [EMDB 1557](#), [EMDB 1558](#), [EMDB 1559](#), [EMDB 1560](#), [EMDB 1561](#), [EMDB 1562](#), [EMDB 1563](#), [EMDB 1564](#), [EMDB 1565](#), [EMDB 1566](#), [EMDB 1567](#), [EMDB 1568](#), [EMDB 1569](#), [EMDB 1570](#), [EMDB 1571](#), [EMDB 1572](#), [EMDB 1573](#), [EMDB 1574](#), [EMDB 1575](#), [EMDB 1576](#), [EMDB 1577](#), [EMDB 1578](#), [EMDB 1579](#), [EMDB 1580](#), [EMDB 1581](#), [EMDB 1582](#), [EMDB 1583](#), [EMDB 1584](#), [EMDB 1585](#), [EMDB 1586](#), [EMDB 1587](#), [EMDB 1588](#), [EMDB 1589](#), [EMDB 1590](#), [EMDB 1591](#), [EMDB 1592](#), [EMDB 1593](#), [EMDB 1594](#), [EMDB 1595](#), [EMDB 1596](#), [EMDB 1597](#), [EMDB 1598](#), [EMDB 1599](#), [EMDB 1600](#), [EMDB 1601](#), [EMDB 1602](#), [EMDB 1603](#), [EMDB 1604](#), [EMDB 1605](#), [EMDB 1606](#), [EMDB 1607](#), [EMDB 1608](#), [EMDB 1609](#), [EMDB 1610](#), [EMDB 1611](#), [EMDB 1612](#), [EMDB 1613](#), [EMDB 1614](#), [EMDB 1615](#), [EMDB 1616](#), [EMDB 1617](#), [EMDB 1618](#), [EMDB 1619](#), [EMDB 1620](#), [EMDB 1621](#), [EMDB 1622](#), [EMDB 1623](#), [EMDB 1624](#), [EMDB 1625](#), [EMDB 1626](#), [EMDB 1627](#), [EMDB 1628](#), [EMDB 1629](#), [EMDB 1630](#), [EMDB 1631](#), [EMDB 1632](#), [EMDB 1633](#), [EMDB 1634](#), [EMDB 1635](#), [EMDB 1636](#), [EMDB 1637](#), [EMDB 1638](#), [EMDB 1639](#), [EMDB 1640](#), [EMDB 1641](#), [EMDB 1642](#), [EMDB 1643](#), [EMDB 1644](#), [EMDB 1645](#), [EMDB 1646](#), [EMDB 1647](#), [EMDB 1648](#), [EMDB 1649](#), [EMDB 1650](#), [EMDB 1651](#), [EMDB 1652](#), [EMDB 1653](#), [EMDB 1654](#), [EMDB 1655](#), [EMDB 1656](#), [EMDB 1657](#), [EMDB 1658](#), [EMDB 1659](#), [EMDB 1660](#), [EMDB 1661](#), [EMDB 1662](#), [EMDB 1663](#), [EMDB 1664](#), [EMDB 1665](#), [EMDB 1666](#), [EMDB 1667](#), [EMDB 1668](#), [EMDB 1669](#), [EMDB 1670](#), [EMDB 1671](#), [EMDB 1672](#), [EMDB 1673](#), [EMDB 1674](#), [EMDB 1675](#), [EMDB 1676](#), [EMDB 1677](#), [EMDB 1678](#), [EMDB 1679](#), [EMDB 1680](#), [EMDB 1681](#), [EMDB 1682](#), [EMDB 1683](#), [EMDB 1684](#), [EMDB 1685](#), [EMDB 1686](#), [EMDB 1687](#), [EMDB 1688](#), [EMDB 1689](#), [EMDB 1690](#), [EMDB 1691](#), [EMDB 1692](#), [EMDB 1693](#), [EMDB 1694](#), [EMDB 1695](#), [EMDB 1696](#), [EMDB 1697](#), [EMDB 1698](#), [EMDB 1699](#), [EMDB 1700](#), [EMDB 1701](#), [EMDB 1702](#), [EMDB 1703](#), [EMDB 1704](#), [EMDB 1705](#), [EMDB 1706](#), [EMDB 1707](#), [EMDB 1708](#), [EMDB 1709](#), [EMDB 1710](#), [EMDB 1711](#), [EMDB 1712](#), [EMDB 1713](#), [EMDB 1714](#), [EMDB 1715](#), [EMDB 1716](#), [EMDB 1717](#), [EMDB 1718](#), [EMDB 1719](#), [EMDB 1720](#), [EMDB 1721](#), [EMDB 1722](#), [EMDB 1723](#), [EMDB 1724](#), [EMDB 1725](#), [EMDB 1726](#), [EMDB 1727](#), [EMDB 1728](#), [EMDB 1729](#), [EMDB 1730](#), [EMDB 1731](#), [EMDB 1732](#), [EMDB 1733](#), [EMDB 1734](#), [EMDB 1735](#), [EMDB 1736](#), [EMDB 1737](#), [EMDB 1738](#), [EMDB 1739](#), [EMDB 1740](#), [EMDB 1741](#), [EMDB 1742](#), [EMDB 1743](#), [EMDB 1744](#), [EMDB 1745](#), [EMDB 1746](#), [EMDB 1747](#), [EMDB 1748](#), [EMDB 1749](#), [EMDB 1750](#), [EMDB 1751](#), [EMDB 1752](#), [EMDB 1753](#), [EMDB 1754](#), [EMDB 1755](#), [EMDB 1756](#), [EMDB 1757](#), [EMDB 1758](#), [EMDB 1759](#), [EMDB 1760](#), [EMDB 1761](#), [EMDB 1762](#), [EMDB 1763](#), [EMDB 1764](#), [EMDB 1765](#), [EMDB 1766](#), [EMDB 1767](#), [EMDB 1768](#), [EMDB 1769](#), [EMDB 1770](#), [EMDB 1771](#), [EMDB 1772](#), [EMDB 1773](#), [EMDB 1774](#), [EMDB 1775](#), [EMDB 1776](#), [EMDB 1777](#), [EMDB 1778](#), [EMDB 1779](#), [EMDB 1780](#), [EMDB 1781](#), [EMDB 1782](#), [EMDB 1783](#), [EMDB 1784](#), [EMDB 1785](#), [EMDB 1786](#), [EMDB 1787](#), [EMDB 1788](#), [EMDB 1789](#), [EMDB 1790](#), [EMDB 1791](#), [EMDB 1792](#), [EMDB 1793](#), [EMDB 1794](#), [EMDB 1795](#), [EMDB 1796](#), [EMDB 1797](#), [EMDB 1798](#), [EMDB 1799](#), [EMDB 1800](#), [EMDB 1801](#), [EMDB 1802](#), [EMDB 1803](#), [EMDB 1804](#), [EMDB 1805](#), [EMDB 1806](#), [EMDB 1807](#), [EMDB 1808](#), [EMDB 1809](#), [EMDB 1810](#), [EMDB 1811](#), [EMDB 1812](#), [EMDB 1813](#), [EMDB 1814](#), [EMDB 1815](#), [EMDB 1816](#), [EMDB 1817](#), [EMDB 1818](#), [EMDB 1819](#), [EMDB 1820](#), [EMDB 1821](#), [EMDB 1822](#), [EMDB 1823](#), [EMDB 1824](#), [EMDB 1825](#), [EMDB 1826](#), [EMDB 1827](#), [EMDB 1828](#), [EMDB 1829](#), [EMDB 1830](#), [EMDB 1831](#), [EMDB 1832](#), [EMDB 1833](#), [EMDB 1834](#), [EMDB 1835](#), [EMDB 1836](#), [EMDB 1837](#), [EMDB 1838](#), [EMDB 1839](#), [EMDB 1840](#), [EMDB 1841](#), [EMDB 1842](#), [EMDB 1843](#), [EMDB 1844](#), [EMDB 1845](#), [EMDB 1846](#), [EMDB 1847](#), [EMDB 1848](#), [EMDB 1849](#), [EMDB 1850](#), [EMDB 1851](#), [EMDB 1852](#), [EMDB 1853](#), [EMDB 1854](#), [EMDB 1855](#), [EMDB 1856](#), [EMDB 1857](#), [EMDB 1858](#), [EMDB 1859](#), [EMDB 1860](#), [EMDB 1861](#), [EMDB 1862](#), [EMDB 1863](#), [EMDB 1864](#), [EMDB 1865](#), [EMDB 1866](#), [EMDB 1867](#), [EMDB 1868](#), [EMDB 1869](#), [EMDB 1870](#), [EMDB 1871](#), [EMDB 1872](#), [EMDB 1873](#), [EMDB 1874](#), [EMDB 1875](#), [EMDB 1876](#), [EMDB 1877](#), [EMDB 1878](#), [EMDB 1879](#), [EMDB 1880](#), [EMDB 1881](#), [EMDB 1882](#), [EMDB 1883](#), [EMDB 1884](#), [EMDB 1885](#), [EMDB 1886](#), [EMDB 1887](#), [EMDB 1888](#), [EMDB 1889](#), [EMDB 1890](#), [EMDB 1891](#), [EMDB 1892](#), [EMDB 1893](#), [EMDB 1894](#), [EMDB 1895](#), [EMDB 1896](#), [EMDB 1897](#), [EMDB 1898](#), [EMDB 1899](#), [EMDB 1900](#), [EMDB 1901](#), [EMDB 1902](#), [EMDB 1903](#), [EMDB 1904](#), [EMDB 1905](#), [EMDB 1906](#), [EMDB 1907](#), [EMDB 1908](#), [EMDB 1909](#), [EMDB 1910](#), [EMDB 1911](#), [EMDB 1912](#), [EMDB 1913](#), [EMDB 1914](#), [EMDB 1915](#), [EMDB 1916](#), [EMDB 1917](#), [EMDB 1918](#), [EMDB 1919](#), [EMDB 1920](#), [EMDB 1921](#), [EMDB 1922](#), [EMDB 1923](#), [EMDB 1924](#), [EMDB 1925](#), [EMDB 1926](#), [EMDB 1927](#), [EMDB 1928](#), [EMDB 1929](#), [EMDB 1930](#), [EMDB 1931](#), [EMDB 1932](#), [EMDB 1933](#), [EMDB 1934](#), [EMDB 1935](#), [EMDB 1936](#), [EMDB 1937](#), [EMDB 1938](#), [EMDB 1939](#), [EMDB 1940](#), [EMDB 1941](#), [EMDB 1942](#), [EMDB 1943](#), [EMDB 1944](#), [EMDB 1945](#), [EMDB 1946](#), [EMDB 1947](#), [EMDB 1948](#), [EMDB 1949](#), [EMDB 1950](#), [EMDB 1951](#), [EMDB 1952](#), [EMDB 1953](#), [EMDB 1954](#), [EMDB 1955](#), [EMDB 1956](#), [EMDB 1957](#), [EMDB 1958](#), [EMDB 1959](#), [EMDB 1960](#), [EMDB 1961](#), [EMDB 1962](#), [EMDB 1963](#), [EMDB 1964](#), [EMDB 1965](#), [EMDB 1966](#), [EMDB 1967](#), [EMDB 1968](#), [EMDB 1969](#), [EMDB 1970](#), [EMDB 1971](#), [EMDB 1972](#), [EMDB 1973](#), [EMDB 1974](#), [EMDB 1975](#), [EMDB 1976](#), [EMDB 1977](#), [EMDB 1978](#), [EMDB 1979](#), [EMDB 1980](#), [EMDB 1981](#), [EMDB 1982](#), [EMDB 1983](#), [EMDB 1984](#), [EMDB 1985](#), [EMDB 1986](#), [EMDB 1987](#), [EMDB 1988](#), [EMDB 1989](#), [EMDB 1990](#), [EMDB 1991](#), [EMDB 1992](#), [EMDB 1993](#), [EMDB 1994](#), [EMDB 1995](#), [EMDB 1996](#), [EMDB 1997](#), [EMDB 1998](#), [EMDB 1999](#), [EMDB 2000](#), [EMDB 2001](#), [EMDB 2002](#), [EMDB 2003](#), [EMDB 2004](#), [EMDB 2005](#), [EMDB 2006](#), [EMDB 2007](#), [EMDB 2008](#), [EMDB 2009](#), [EMDB 2010](#), [EMDB 2011](#), [EMDB 2012](#), [EMDB 2013](#), [EMDB 2014](#), [EMDB 2015](#), [EMDB 2016](#), [EMDB 2017](#), [EMDB 2018](#), [EMDB 2019](#), [EMDB 2020](#), [EMDB 2021](#), [EMDB 2022](#), [EMDB 2023](#), [EMDB 2024](#), [EMDB 2025](#), [EMDB 2026](#), [EMDB 2027](#), [EMDB 2028](#), [EMDB 2029](#), [EMDB 2030](#), [EMDB 2031](#), [EMDB 2032](#), [EMDB 2033](#), [EMDB 2034](#), [EMDB 2035](#), [EMDB 2036](#), [EMDB 2037](#), [EMDB 2038](#), [EMDB 2039](#), [EMDB 2040](#), [EMDB 2041](#), [EMDB 2042](#), [EMDB 2043](#), [EMDB 2044](#), [EMDB 2045](#), [EMDB 2046](#), [EMDB 2047](#), [EMDB 2048](#), [EMDB 2049](#), [EMDB 2050](#), [EMDB 2051](#), [EMDB 2052](#), [EMDB 2053](#), [EMDB 2054](#), [EMDB 2055](#), [EMDB 2056](#), [EMDB 2057](#), [EMDB 2058](#), [EMDB 2059](#), [EMDB 2060](#), [EMDB 2061](#), [EMDB 2062](#), [EMDB 2063](#), [EMDB 2064](#), [EMDB 2065](#), [EMDB 2066](#), [EMDB 2067](#), [EMDB 2068](#), [EMDB 2069](#), [EMDB 2070](#), [EMDB 2071](#), [EMDB 2072](#), [EMDB 2073](#), [EMDB 2074](#), [EMDB 2075](#), [EMDB 2076](#), [EMDB 2077](#), [EMDB 2078](#), [EMDB 2079](#), [EMDB 2080](#), [EMDB 2081](#), [EMDB 2082](#), [EMDB 2083](#), [EMDB 2084](#), [EMDB 2085](#), [EMDB 2086](#), [EMDB 2087](#), [EMDB 2088](#), [EMDB 2089](#), [EMDB 2090](#), [EMDB 2091](#), [EMDB 2092](#), [EMDB 2093](#), [EMDB 2094](#), [EMDB 2095](#), [EMDB 2096](#), [EMDB 2097](#), [EMDB 2098](#), [EMDB 2099](#), [EMDB 2100](#), [EMDB 2101](#), [EMDB 2102](#), [EMDB 2103](#), [EMDB 2104](#), [EMDB 2105](#), [EMDB 2106](#), [EMDB 2107](#), [EMDB 2108](#), [EMDB 2109](#), [EMDB 2110](#), [EMDB 2111](#), [EMDB 2112](#), [EMDB 2113](#), [EMDB 2114](#), [EMDB 2115](#), [EMDB 2116](#), [EMDB 2117](#), [EMDB 2118](#), [EMDB 2119](#), [EMDB 2120](#), [EMDB 2121](#), [EMDB 2122](#), [EMDB 2123](#), [EMDB 2124](#), [EMDB 2125](#), [EMDB 2126](#), [EMDB 2127](#), [EMDB 2128](#), [EMDB 2129](#), [EMDB 2130](#), [EMDB 2131](#), [EMDB 2132](#), [EMDB 2133](#), [EMDB 2134](#), [EMDB 2135](#), [EMDB 2136](#), [EMDB 2137](#), [EMDB 2138](#), [EMDB 2139](#), [EMDB 2140](#), [EMDB 2141](#), [EMDB 2142](#), [EMDB 2143](#), [EMDB 2144](#), [EMDB 2145](#), [EMDB 2146](#), [EMDB 2147](#), [EMDB 2148](#), [EMDB 2149](#), [EMDB 2150](#), [EMDB 2151](#), [EMDB 2152](#), [EMDB 2153](#), [EMDB 2154](#), [EMDB 2155](#), [EMDB 2156](#), [EMDB 2157](#), [EMDB 2158](#), [EMDB 2159](#), [EMDB 2160](#), [EMDB 2161](#), [EMDB 2162](#), [EMDB 2163](#), [EMDB 2164](#), [EMDB 2165](#), [EMDB 2166](#), [EMDB 2167](#), [EMDB 2168](#), [EMDB 2169](#), [EMDB 2170](#), [EMDB 2171](#), [EMDB 2172](#), [EMDB 2173](#), [EMDB 2174](#), [EMDB 2175](#), [EMDB 2176](#), [EMDB 2177](#), [EMDB 2178](#), [EMDB 2179](#), [EMDB 2180](#), [EMDB 2181](#), [EMDB 2182](#), [EMDB 2183](#), [EMDB 2184](#), [EMDB 2185](#), [EMDB 2186](#), [EMDB 2187](#), [EMDB 2188](#), [EMDB 2189](#), [EMDB 2190](#), [EMDB 2191](#), [EMDB 2192](#), [EMDB 2193](#), [EMDB 2194](#), [EMDB 2195](#), [EMDB 2196](#), [EMDB 2197](#), [EMDB 2198](#), [EMDB 2199](#), [EMDB 2200](#), [EMDB 2201](#), [EMDB 2202](#), [EMDB 2203](#), [EMDB 2204](#), [EMDB 2205](#), [EMDB 2206](#), [EMDB 2207](#), [EMDB 2208](#), [EMDB 2209](#), [EMDB 2210](#), [EMDB 2211](#), [EMDB 2212](#), [EMDB 2213](#), [EMDB 2214](#), [EMDB 2215](#), [EMDB 2216](#), [EMDB 2217](#), [EMDB 2218](#), [EMDB 2219](#), [EMDB 2220](#), [EMDB 2221](#), [EMDB 2222](#), [EMDB 2223](#), [EMDB 2224](#), [EMDB 2225](#), [EMDB 2226](#), [EMDB 2227](#), [EMDB 2228](#), [EMDB 2229](#), [EMDB 2230](#), [EMDB 2231](#), [EMDB 2232](#), [EMDB 2233](#), [EMDB 2234](#), [EMDB 2235](#), [EMDB 2236](#), [EMDB 2237](#), [EMDB 2238](#), [EMDB 2239](#), [EMDB 2240](#), [EMDB 2241](#), [EMDB 2242](#), [EMDB 2243](#), [EMDB 2244](#), [EMDB 2245](#), [EMDB 2246](#), [EMDB 2247](#), [EMDB 2248](#), [EMDB 2249](#), [EMDB 2250](#), [EMDB 2251](#), [EMDB 2252](#), [EMDB 2253](#), [EMDB 2254](#), [EMDB 2255](#), [EMDB 2256](#), [EMDB 2257](#), [EMDB 2258](#), [EMDB 2259](#), [EMDB 2260](#), [EMDB 2261](#), [EMDB 2262](#), [EMDB 2263](#), [EMDB 2264](#), [EMDB 2265](#), [EMDB 2266](#), [EMDB 2267](#), [EMDB 2268](#), [EMDB 2269](#), [EMDB 2270](#), [EMDB 2271](#), [EMDB 2272](#), [EMDB 2273](#), [EMDB 2274](#), [EMDB 2275](#), [EMDB 2276](#), [EMDB 2277](#), [EMDB 2278](#), [EMDB 2279](#), [EMDB 2280](#), [EMDB 2281](#), [EMDB 2282](#), [EMDB 2283](#), [EMDB 2284](#), [EMDB 2285](#), [EMDB 2286](#), [EMDB 2287](#), [EMDB 2288](#), [EMDB 2289](#), [EMDB 2290](#), [EMDB 2291](#), [EMDB 2292](#), [EMDB 2293](#), [EMDB 2294](#), [EMDB 2295](#), [EMDB 2296](#), [EMDB 2297](#), [EMDB 2298](#), [EMDB 2299](#), [EMDB 2300](#), [EMDB 2301](#), [EMDB 2302](#), [EMDB 2303](#), [EMDB 2304](#), [EMDB 2305](#), [EMDB 2306](#), [EMDB 2307](#), [EMDB 2308](#), [EMDB 2309](#), [EMDB 2310](#), [EMDB 2311](#), [EMDB 2312](#), [EMDB 2313](#), [EMDB 2314](#), [EMDB 2315](#), [EMDB 2316](#), [EMDB 2317](#), [EMDB 2318](#), [EMDB 2319](#), [EMDB 2320](#), [EMDB 2321](#), [EMDB 2322](#), [EMDB 2323](#), [EMDB 2324](#), [EMDB 2325](#), [EMDB 2326](#), [EMDB 2327](#), [EMDB 2328](#), [EMDB 2329](#), [EMDB 2330](#), [EMDB 2331](#), [EMDB 2332](#), [EMDB 2333](#), [EMDB 2334](#), [EMDB 2335](#), [EMDB 2336](#), [EMDB 2337](#), [EMDB 2338](#), [EMDB 2339](#), [EMDB 2340](#), [EMDB 2341](#), [EMDB 2342](#), [EMDB 2343](#), [EMDB 2344](#), [EMDB 2345](#), [EMDB 2346](#), [EMDB 2347](#), [EMDB 2348](#), [EMDB 2349](#), [EMDB 2350](#), [EMDB 2351](#), [EMDB 2352](#), [EMDB 2353](#), [EMDB 2354](#), [EMDB 235](#)



# Is it a practical suggestion?

- Google, Yahoo! and Bing decided to use microdata to show search results more valuable.
- Some vocabularies have already applied to search results.
- E.g.

## [Beef Stew Scented with Horseradish Recipe ... - Food Network](#)



[www.foodnetwork.com](#) > [Recipes](#) > [Beef](#) - [Cached](#)

★★★★★ Rating: 4 - 2 reviews - 2 hrs 30 mins

Mar 7, 2012 – Get this all-star, easy-to-follow Food Network **Beef Stew Scented with Horseradish** recipe from Rachael Ray.

## [Review: Samsung Galaxy Note II is a solid jumbo smartphone \[Video\]](#)

[articles.latimes.com/.../la-fi-tn-samsung-galaxy-note-...](#) - [Cached](#)



by Salvador Rodriguez - in 409 Google+ circles

Dec 8, 2012 – If you're an NFL lineman or a basketball player, **Samsung** has got the phone for you. The South Korean company's **Galaxy Note II** is a massive ...



# Schema.org

- Provide a collection of schemas (html tags)
- Bing, Google, Yahoo! and Yandex rely on this markup to improve the display of search results, making it easier for people to find the right web pages. (quoted by schema.org)
- We proposed “schema.org” extensions for “**BiologicalDatabaseEntry**” and “**BiologicalDatabase**”.
- Schema.org proposals : <http://www.w3.org/wiki/WebSchemas/SchemaDotOrgProposals>



# Properties for BiologicalDatabaseEntry

Well targeted

entryID	additionalType	dateCreated
isEntryof	description	dateModified
taxon	image	keywords
seeAlso	url	provider
reference	alternativeHeadline	breadcrumb
name	inLanguage	



## Related Link for our proposal

- WebSchemas proposal 'Biological Databases' for schema.org
  - <http://www.w3.org/wiki/WebSchemas/BioDatabases>
- Discussions at BioHackathon
  - <https://github.com/dbcls/bh12/wiki/Schema.org-extension>
- Discussions at BH12.12 (Japanese only)
  - <http://wiki.lifesciencedb.jp/mw/index.php/BH12.12/schema.org>



# How to markup ?

Declaration

```
<div itemscope itemtype="http://schema.org/BiologicalDatabaseEntry">
```

ID

```
<span itemprop="entryID">1556</span>
```

Specied

```
<span itemprop="taxon" itemscope itemtype="http://schema.org/BiologicalDatabaseEntry">
```

```
<span itemprop="name">Bacillus subtilis</span>
```

```
</span>
```

Deposition:

```
<span itemprop="dateCreated">2008-09-08</span>
```

Last update:

```
<span itemprop="dateModified">2012-10-24</span>
```

```
</div>
```

Specify Property and  
markup with normal tag

## ▪ Crawl these microdata

```
@uri=http://cellbank.nibio.go.jp/~cellbank/cgi-bin/search_res_det.cgi?DB_NM=1&ID=1&32
@title=PB-3c - JCRB細胞バンク
@db=cellbank
@entryID=IFO50267
@species=Mouse
@taxonID=10090
@BiologicalDatabaseEntry_ResourceAvailability=F
@BiologicalDatabaseEntry_entryID=IFO50267
@BiologicalDatabaseEntry_name=PB-3c
@BiologicalDatabaseEntry_description=IL-3-dependent mouse bone marrow cell line.
@BiologicalDatabaseEntry_dateCreated=1990-10-31
@BiologicalDatabaseEntry_keywords=mouse
@BiologicalDatabaseEntry_OriginalTissue=bone marrow
@BiologicalDatabase_url=http://cellbank.nibio.go.jp/
@BiologicalDatabase_name=JCRB Cell Bank
@Organization_url=http://www.nibio.go.jp/
@Organization_name=National Institute of Biomedical Innovation (NIBIO)

IFO50267 PB-3c マウス 血球・リンパ系 肥満(マスト)細胞株 IL-3-dependent mouse bone marrow cell line.
mouse Mus musculus not done bone marrow infinite lymphoblast-like mast cell line tumor
Ball, P. E. Shimizu, T. 1990 RPMI1640 with 10% fetal bovine serum, 10% WEHI-3b conditioned medium
and 50 uM 2-mercaptoethanol dilution 5 % registered bone marrow
```



## • Reflect Search Results

### PB-3c - JCRB細胞バンク

<http://cellbank.nibio.go.jp/> JCRB細胞バンク

Biological Resources | Human, Animals (non-human) | Cell & Tissue

ID: IFO50267 - Species: Mouse - Original Tissue: bone marrow

IFO50267 PB-3c マウス 血球・リンパ系 肥満(マスト)細胞株 IL-3-dependent mouse bone marrow cell line. mouse Mus musculus not done bone marrow infinite lymphoblast-like mast cell line tu ....



- Pop-up image of species

**PB-3c - JCRB細胞バンク**  
<http://cellbank.nibio.go.jp/> JCRB細胞バンク  
Biological Resources | Human, Animals (non-human) | Cell & Tissue  
ID: IFO50267 - Species: Mouse - Original Tissue: bone marrow  
IFO50267 PB-3c マウス 血液・リンパ系 配適 Mus musculus -3-dependent mouse bone marrow cell line. mouse Mus musculus not bone bone marrow infinite lymphoblast-like mast cell line tu .....

Mus musculus



- Refine by ID



Search Results: 1 hits      Option: Synonym Expansion + (Default)

**PB-3c - JCRB細胞バンク**  
<http://cellbank.nibio.go.jp/> JCRB細胞バンク  
Biological Resources | Human, Animals (non-human) | Cell & Tissue  
ID: IFO50267 - Species: Mouse - Original Tissue: bone marrow



# Sagace support

KEGG DISEASE: H00653

<http://www.kegg.jp/>

KEGG DISEASE (Diseases viewed as perturbed states)

その他のデータベース | ヒト | その他

- ID: H00653 - Species: Human - Disease: マルファン症候群

Entry H00653 Disease Name **marfan syndrome**, including

Neonatal MFS; Atypically severe MFS; New variant of MFS

(MFS) is a relatively common autosomal dominant di ....

ICD-10: Q87.4 - OMIM: 154700

Image

- Biological DatabaseEntry
  - entryID, taxon , dateModified, seeAlso
- MedicalCode
  - Disease (codeValue, codingSystem)
- Testing
  - (Resourceavailability, originalTissue)



# Future perspective

- Improve search results
  - 1.Reference
- Not only for search results.
  - 1.We can create interesting functions.
  - 2.Developers of databases had better to feel benefit immediately.

## Idea 1 : Reference

### KEGG DISEASE: H00653

<http://www.kegg.jp/>

KEGG DISEASE (Diseases viewed as perturbed states of the molecular system)

その他のデータベース | ヒト | その他

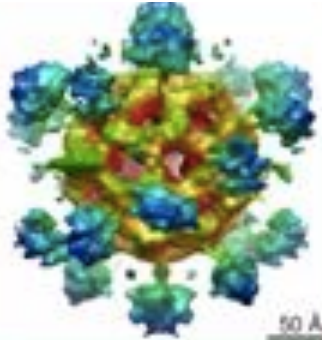
- ID: H00653 - Species: Human - Disease: Marfan syndrome - Reference : 4

Entry **H00653** disease Name Marfan syndrome, including: Marfan syndrome (MFS):

Neonatal MFS; Aty  
connective tissue.

- Robinson PN,Booms P., The molecular pathogenesis of the Marfan syndrome., Cell Mol Life Sci, 2001 Oct, 58(11):1698-707, [PubMed:11706995](#)
- Robinson PN,Arteaga-Solis E,Baldock C,Collod-Béroud G,Booms P,De Paepe A,Dietz HC,Guo G,Handford PA,Judge DP,Kielty CM,Loeys B,Milewicz DM,Ney A,Ramirez F,Reinhardt DP,Tiedemann K,Whiteman P,Godfrey M., The molecular genetics of Marfan syndrome and related disorders., J Med Genet, 2006 Oct, 43(10):769-87, [PubMed:16571647](#)
- Robinson PN,Booms P,Katzke S,Ladewig M,Neumann L,Palz M,Pregla R,Tiecke F,Rosenberg T., Mutations of FBN1 and genotype-phenotype correlations in Marfan syndrome and related fibrillinopathies., Hum Mutat, 2002 Sep, 20(3):153-61, [PubMed:12203987](#)
- Robinson PN,Godfrey M., The molecular genetics of Marfan syndrome and related microfibrilopathies., J Med Genet, 2000 Jan, 37(1):9-25, [PubMed:10633129](#)

# Idea2: Developers can feel benefit immediately (By applying Sagace function)



50 Å

表: 個別 すべて

- エントリ情報
- 文献
- マップデータ
- 添付情報
- 試料
- 実験
- 解析
- ダウンロード

文献

引用 - Primary

文献	<a href="#">Science, Vol. 322, Issue 5898, Page 92-6, Year 2008</a>
タイトル	Molecular architecture of the "stressosome," a signal integration and transduction hub.
著者・データ登録者	<a href="#">Jon Marles-Wright</a> , <a href="#">Tim Grant</a> , <a href="#">Olivier Delumeau</a> , <a href="#">Gijs van Duinen</a> , <a href="#">Susan J Firbank</a> , <a href="#">Peter J Lewis</a> , <a href="#">James W Murray</a> , <a href="#">Joseph A Newman</a> , <a href="#">Maureen B Quin</a> , <a href="#">Paul R Race</a> , <a href="#">Alexis Rohou</a> , <a href="#">Willem Tichelaar</a> , <a href="#">Marin van Heel</a> , <a href="#">Richard J Lewis</a>
所属	Institute for Cell and Molecular Biosciences, Newcastle University, Newcastle-upon-Tyne NE2 4HH, UK.
試薬・成分	Bacterial Proteins, Multiprotein Complexes, Phosphoproteins, RsbR protein, <i>Bacillus subtilis</i> , Sigma Factor, RsbT protein, <i>Bacillus subtilis</i> (EC 2.7.1.-), Protein-Serine-Threonine Kinases (EC 2.7.1.1)
キーワード	<a href="#">Amino Acid Sequence</a> , <a href="#">Bacterial Proteins</a> , <a href="#">Cryo-electron Microscopy</a> , <a href="#">X-Ray</a> , <a href="#">Image Processing, Computer Graphics</a> , <a href="#">Biological</a> , <a href="#">Models, Molecular</a> , <a href="#">Data</a> , <a href="#">Multiprotein Complexes</a> , <a href="#">Phosphorylation</a> , <a href="#">Protein Structure</a>

Bacillus subtilis

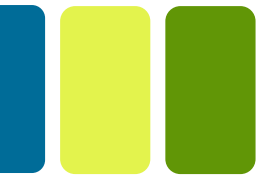




# Ask for your help

- We'd like to finalize our proposal!
- Please markup your own site or database and give me feedback.
- Could you think exciting functions with us?





Thank you for  
listening!



# Taxon (species)

```
<span itemprop="taxon" itemscope itemtype="http://schema.org/BiologicalDatabaseEntry">  
Organism: <span itemprop="name">Homo sapiens</span> (human)  
Taxonomy ID: <a itemprop="url" href="http://www.uniprot.org/taxonomy/9606">  
    <span itemprop="entryID">9606</span></a>  
</span>
```

- It can use many kinds of taxonomic ID systems
- E.g. NCBI, uniprot



## seeAlso

- KEGG:  
`<span itemprop="seeAlso" itemscope itemtype="http://  
schema.org/BiologicalDatabaseEntry">  
<a itemprop="url" href="http://purl.uniprot.org/kegg/hsa:  
353174">  
<span itemprop="name">hsa:353174</span>  
</a>  
</span>`
- We decided to use not only URL but also using other tags in biological database entry, so we applied seeAlso not relatedLink.



# References

- References:

Kanehisa M, et al. Nucleic Acids Res. 40, D109-D114 (2012).  
[<meta itemprop='reference' content='pmid:22080510' />  
<a href="http://www.ncbi.nlm.nih.gov/pubmed/  
22080510">pubmed</a>]

We applied reference because we can specify reference systems by using content tag.